Reduced Degree-of-Freedom STAP with Knowledge-Aided Data Pre-Whitening

Christopher M. Teixeira, Jameson S. Bergin, Paul M. Techau

Information Systems Laboratories, Inc.
8130 Boone Blvd. ste. 500
Vienna, VA 22182 USA
cmt@isl-inc.com, jsb@isl-inc.com, pmt@isl-inc.com

ABSTRACT

A major thrust of DARPA's Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER) program is to develop radar signal processing algorithms that exploit the ever expanding body of a priori knowledge about the sensor operating environment. Knowledge sources include digital terrain maps, land coverage data, and the locations of man-made features such as roads and buildings. This paper presents an extension of knowledge-aided data pre-whitening techniques based on colored diagonal loading [1,2] to low-rank space-time adaptive processing (STAP) implementations such as extended factored post-Doppler STAP [3] ('multi-bin post-Doppler' STAP). The new implementations maintain the same desirable property of "blending" the information contained in the observed radar data and the a priori knowledge sources but offer added flexibility in reducing the computational complexity of the overall space-time beamforming solution. Furthermore, the approach has the potential benefit of minimizing sample support requirements by reducing the interference rank within the chosen low-rank subspace (e.g., post-Doppler) due to the effective pre-whitening of the data. A computationally efficient data domain implementation of the algorithms is developed along with analytical low-rank representations of the a priori ground clutter covariance models (i.e., the colored loading matrix). The performance of these reduced-DoF knowledge-aided beamforming techniques is demonstrated using high-fidelity radar simulation data.

Presented at the 2003 KASSPER Workshop, Las Vegas, NV, 15-17 April, 2003.

This work was sponsored under Air Force Contract F30602-02-C-0005.

Approved for Public Release. Distribution Unlimited.

1. INTRODUCTION

DARPA's Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER) program is developing radar signal processing algorithms that exploit *a priori* knowledge about the sensor operating environment. Knowledge sources include digital terrain maps, land coverage data, and the locations of man-made features such as roads and buildings. Since radar clutter is highly dependent on the various features represented by these knowledge sources (e.g., clutter power is a strong function of terrain height and slope) it is logical to believe that exploiting them will improve radar performance.

The main problem being addressed under KASSPER is that of incorporating knowledge sources in the beamformer in an attempt to improve the performance of GMTI radar. A primary focus of the program is to demonstrate that by using *a priori* information we can significantly reduce sample support requirements of the adaptive clutter filters. This will generally result in improved performance in heterogeneous clutter environments caused by effects such as site-specific terrain [4], internal clutter motion (ICM) [5], and high densities of targets [6,7].

Knowledge sources can generally be used to influence the beamformer response in either a *direct* or *indirect* manner. The indirect uses would include procedures including the selection of training data regions based on terrain feature databases (e.g., [6]) such as the National Imagery and Mapping Agency (NIMA) digital terrain elevation data (DTED) and digital features analysis data (DFAD). Direct uses would include explicitly forcing nulls in the beamformer response pattern based on the known location of interference sources. This paper presents an approach for incorporating knowledge sources *directly* in the space-time beamformer of airborne adaptive radars to improve the cancellation of ground clutter. The *a priori* knowledge is incorporated into the beamformer response via constraints involving an *a priori* site-specific clutter covariance model. The structure of this model is similar to the models described in [4-6] and is given by:

$$\mathbf{R}_c = \sum_{p=1}^{P_c} |\alpha_p|^2 \mathbf{v}(\theta_p, f_p)\mathbf{v}^H(\theta_p, f_p) \circ \mathbf{T}_p \quad (1)$$

where α_p, θ_p, and f_p are the complex amplitude, direction of arrival (DoA), and Doppler shift, respectively, for the p^{th} ground clutter patch, $\mathbf{v}(\theta, f)$ is the space-time response vector for a signal with DoA θ and Doppler shift f, $\circ$ is the Hadamard (element-wise) product, and $\mathbf{T}_p$ is a covariance matrix taper [8] that allows for incorporation of effects such as calibration errors and ICM (e.g., [5]). If this matrix is known perfectly then we would expect optimal beamformer performance. Of course, this matrix is never known perfectly in practical applications due to various real-world effects. However, parts of the model may be known and can potentially be used in the beamformer solution. The techniques developed in this paper attempt to minimize the amount of *a priori* knowledge required for computing $\mathbf{R}_c$ so as to make the solution more useful in practice where availability of databases, sample support, and computational resources are likely to be limited.

Section 2 develops the knowledge-aided constraints and appropriate weight vectors for both full- and reduced-degree of freedom (DoF) STAP. Section 3 discusses techniques for efficient implementation of the developed algorithms. Section 4 presents results using high-fidelity simulated radar data and Section 5 provides a summary and conclusions.

2. KNOWLEDGE-AIDED CONSTRAINTS

We begin by presenting the following space-time (e.g., elements and pulses) interference model that helps motivate the algorithms developed in this paper. The interference received on the array is represented by $\mathbf{x} = \mathbf{x}_c \circ \mathbf{t} + \mathbf{n}$, where $\mathbf{x}_c$ represents the radar ground clutter with second order statistics that will be assumed to be known to some degree *a priori*, $\mathbf{t}$ is a vector that represents small unknown random modulations and/or errors on the clutter signal (e.g. ICM, calibration errors, etc.), and $\mathbf{n}$ is the white thermal noise. The modulation $\mathbf{t}$ will typically have the form $\mathbf{t} = \mathbf{1} + \tilde{\mathbf{t}}$ where $\mathbf{1}$ is a vector of ones and $\tilde{\mathbf{t}}$ is a zero-mean random vector with variance that is typically << 1. If we assume that $\tilde{\mathbf{t}}$ and $\mathbf{x}_c$ are uncorrelated then the covariance matrix of $\mathbf{x}$ is given as,

$$\mathbf{R}_{xx} = E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{R}_c + \mathbf{R}_c \circ \tilde{\mathbf{T}} + \sigma^2\mathbf{I} \quad (2)$$

where $\tilde{\mathbf{T}} = E\{\tilde{\mathbf{t}}\tilde{\mathbf{t}}^H\}$ and $\mathbf{I}$ is an identity matrix. We see that the clutter covariance matrix is comprised of a known component, $\mathbf{R}_c$, and an unknown component, $\mathbf{R}_c \circ \tilde{\mathbf{T}}$. Therefore we will be interested in beamforming solutions that combine both deterministic and adaptive filtering to null these components, respectively. We note that the notion of prefiltering the ground clutter [9] or large clutter discretes [10] followed by adaptive processing has been previously suggested.

2.1. Full-DoF STAP

We first review the incorporation of knowledge-aided constraints for full-DoF STAP, as presented in [2]. The optimization problem to be solved is

$$\min_{\mathbf{w}} E\{|\mathbf{w}^H\mathbf{x}|^2\} \quad \text{s.t.} \quad \begin{cases} \mathbf{w}^H\mathbf{v} = 1 \\ \mathbf{w}^H\mathbf{R}_c\mathbf{w} \le \delta_d \\ \mathbf{w}^H\mathbf{w} \le \delta_L \end{cases} \quad (3)$$

where $\mathbf{x}$ is the full-DoF ($[MN \times 1]$, M = number of pulses, N = number of elements) data vector, $\mathbf{v}$ represents a desired steering direction and Doppler shift and δ_L is chosen to give a desired maximum gain on white noise [11]. The quadratic inequality constraint $\mathbf{w}^H\mathbf{R}_c\mathbf{w} \le \delta_d$ [12] incorporates the *a priori* knowledge by "limiting" the solution to have no more than a maximum desired gain on the dominant subspace occupied by the *a priori* covariance matrix $\mathbf{R}_c$, which is computed using available knowledge about the clutter environment. We note that $\mathbf{R}_c$ will typically consist of a dominant subspace that is much smaller in dimension than the full space-time dimension of the system (e.g. Brennan's Rule [13]). The solution to this optimization problem is (e.g., [14]),

$$\mathbf{w} = \frac{(\mathbf{R}_{xx} + \beta_d\mathbf{R}_c + \beta_L\mathbf{I})^{-1}\mathbf{v}}{\mathbf{v}^H(\mathbf{R}_{xx} + \beta_d\mathbf{R}_c + \beta_L\mathbf{I})^{-1}\mathbf{v}} = \frac{(\mathbf{R}_{xx} + \mathbf{Q})^{-1}\mathbf{v}}{\mathbf{v}^H(\mathbf{R}_{xx} + \mathbf{Q})^{-1}\mathbf{v}}. \quad (4)$$

In practice the matrix $\mathbf{R}_{xx}$ will be replaced with an estimate of the data covariance matrix computed using available auxiliary radar data snapshots. We see that the solution results in the usual diagonal loading [11] term $\beta_L\mathbf{I}$ plus a "colored" loading [15,16] term $\beta_d\mathbf{R}_c$. The loading levels, embedded in the matrix $\mathbf{Q}$ defined in (4), are chosen in order to satisfy the two coupled non-linear inequality relations resulting from (3)

$$\begin{gathered} \mathbf{v}^H(\mathbf{R}_{xx} + \mathbf{Q})^{-1}\mathbf{R}_c(\mathbf{R}_{xx} + \mathbf{Q})^{-1}\mathbf{v} \le \delta_d(\mathbf{v}^H(\mathbf{R}_{xx} + \mathbf{Q})^{-1}\mathbf{v})^2 \\ \mathbf{v}^H(\mathbf{R}_{xx} + \mathbf{Q})^{-2}\mathbf{v} \le \delta_L(\mathbf{v}^H(\mathbf{R}_{xx} + \mathbf{Q})^{-1}\mathbf{v})^2. \end{gathered} \quad (5)$$

An analytical solution for the optimal value of the loading coefficients β_L and β_d does not exist. They will typically be set based on reasonable assumptions about the interference environment and available sample support for estimating $\mathbf{R}_{xx}$. While it is clear that a non-trivial weight vector requires that the white noise gain level be $\delta_L > 0$, it is possible to drive the gain on the *a priori* covariance matrix to zero, $\delta_d \to 0$. This results in the orthogonality of the weight vector to the dominant subspace of this matrix. However, the approach to this limit leads to an unbounded increase in the required value of the colored loading level, β_d. This can be demonstrated explicitly by comparing the solution to the above quadratic constraint problem to the solution when $\delta_d = 0$ precisely, which then reduces the second quadratic constraint to a set of multiple linear constraints that can be solved exactly [2]. Consequently, a finite level of colored loading approximates the orthogonality relation. This is a desirable property since due to unknown errors in the clutter model it may be better to reduce the deterministic "null-depth" on the clutter subspace. Furthermore, the colored loading approach, unlike the linear constraint approach that enforces exact orthogonality, has the potential for efficient implementation as discussed in Section 3.

It is interesting to note that the solution given in (4) results in a "blending" of the information contained in the sample covariance matrix and the *a priori* clutter model. Therefore the solution has the desirable property of combining adaptive and deterministic filtering. In fact, the solution will provide beampatterns that are a mix between the fully adaptive pattern, a fully deterministic filter, and the conventional pattern represented by the constraint $\mathbf{v}$. An interesting area for future research will be to develop rules for setting the covariance "blending" factors based on the characteristics of the operating environment (e.g., expected density of targets, terrain type, etc.) derived from auxiliary databases.

Finally, we also note that the beamformer weights in (4) can be re-written to permit interpretation as a two-stage filter where the first stage "whitens" the data vector using the *a priori* covariance model, which is then followed by an adaptive beamformer based on the whitened data [2].

2.2. Reduced-DoF STAP

For many systems it may be impractical to use the full-DoF formulation for reasons of computational complexity and the absence of sufficient sample support for covariance estimation. In such situations, it is desirable to reduce the number of degrees of freedom to a manageable number while minimizing any loss in performance relative to the full-DoF formulation. A common approach [13] is to break the full-DoF problem into a number of smaller problems via the application of an $[MN \times D]$ (with $D<MN$) transformation matrix, $\mathbf{H}_m$, to the data. The resulting reduced-DoF data and steering vectors, both $[D \times 1]$, are determined using

$$\mathbf{x}_m = \mathbf{H}_m^H \mathbf{x} \quad ; \quad \mathbf{v}_m = \mathbf{H}_m^H \mathbf{v} . \tag{6}$$

The reduced-DoF covariance estimation, $\mathbf{R}_m$, is found in the usual way using available auxiliary radar data snapshots, transformed in this manner. The transformation is also applied to the *a priori* clutter covariance $\mathbf{R}_c$ and thermal noise covariance $\mathbf{R}_n$ models to produce reduced-DoF versions of these quantities

$$\mathbf{R}_{c,m} = \mathbf{H}_m^H \mathbf{R}_c \mathbf{H}_m \quad ; \quad \mathbf{R}_{n,m} = \mathbf{H}_m^H \mathbf{R}_n \mathbf{H}_m . \tag{7}$$

In this paper, we focus on the reduced-DoF technique known as extended factored or multi-bin element-space post-Doppler STAP [3]. With this technique, $D = KN$, where the number of elements (i.e. spatial DoFs) are preserved while the temporal DoFs are reduced, $K < M$. The temporal portion of the transformation matrix, $\mathbf{D}_{m,k}$, consists of the M-component discrete fourier transform at K adjacent orthogonal Doppler frequencies with the Doppler bin under test typically the central bin (i.e. K is odd). The full transformation matrix is then defined

$$\mathbf{H}_m = \mathbf{D}_{m,k} \otimes \mathbf{I}_N \tag{8}$$

where the Kronecker tensor product is indicated and we define $\mathbf{I}_Q$ to be the $[Q \times Q]$ identity matrix. With this formulation, and the assumption of uncorrelated thermal noise, the thermal noise covariance, $\mathbf{R}_n = \mathbf{I}_{MN}$, is transformed to $\mathbf{R}_{n,m} = M\,\mathbf{I}_D$.

In a manner similar to that of the full-DoF case, the reduced-DoF covariance model can be incorporated into an optimization problem for the post-Doppler weights, $\mathbf{w}_m$, as a quadratic constraint

$$\min_{\mathbf{w}_m} E\{|\mathbf{w}_m^H \mathbf{x}_m|^2\} \quad \text{s.t.} \quad \begin{cases} \mathbf{w}_m^H \mathbf{v}_m = 1 \\ \mathbf{w}_m^H \mathbf{R}_{c,m} \mathbf{w}_m \le \delta_{d,m} \\ \mathbf{w}_m^H \mathbf{w}_m \le \delta_{L,m} \end{cases} , \tag{9}$$

along with the usual constraints of unity gain on the desired signal, and gain on the white noise (note that the scalar factors that result from the gain on white noise relation have been absorbed into the reduced-DoF gain factor $\delta_{L,m}$). The quadratic constraint on the reduced-DoF *a priori* covariance seeks to make the weights as orthogonal as possible to the dominant subspace of this matrix. Comparing with (3), the form of the reduced-DoF optimization problem is the same as that of the full-

DoF case, except that reduced-DoF, instead of full-DoF, versions of the data and steering vectors, along with *a priori* covariance and thermal noise matrices, are used. As a consequence, the desired weight vector may be derived either directly, or by analogy with the solution of (3), to be

$$\begin{aligned}\mathbf{w}_m &= \frac{(\mathbf{R}_m + \beta_{d,m}\mathbf{R}_{c,m} + \beta_{L,m}\mathbf{I}_D)^{-1}\mathbf{v}_m}{\mathbf{v}_m^H(\mathbf{R}_m + \beta_{d,m}\mathbf{R}_{c,m} + \beta_{L,m}\mathbf{I}_D)^{-1}\mathbf{v}_m} \\ &= \frac{(\mathbf{R}_m + \mathbf{Q}_m)^{-1}\mathbf{v}_m}{\mathbf{v}_m^H(\mathbf{R}_m + \mathbf{Q}_m)^{-1}\mathbf{v}_m}\end{aligned}. \tag{10}$$

As with the full-DoF case, the reduced-DoF solution has both a diagonal and colored loading component. Consequently, this solution can also be interpretated as a "blending" of adaptive and deterministic information, although less sample support is now required for the adaptive portion. The loading levels should be determined from the reduced-DoF analogue of (5) by simply replacing all full-DoF variables by their reduced-DoF counterparts. Finally, the reduced-DoF solution may also be interpretated as a two-stage filter, with an initial deterministic pre-whitening ($\mathbf{Q}_m^{-1/2}$) stage, followed by an adaptive stage.

3. IMPLEMENTATION ISSUES

This section outlines two important implementation issues associated with colored loading. The first is practical considerations about the clutter covariance model that relaxes the fidelity of the *a priori* knowledge required by the technique and the other is methods for efficient implementation of colored loading in the data domain. Finally, we examine the computational complexity of two alternative approaches to implementing reduced-DoF colored loading.

3.1. Practical Clutter Covariance Models

In practice it may not be possible to compute the exact colored loading matrix given in (1) if *a priori* knowledge data such as DTED/DFAD is missing or if computational resources are limited. Therefore, it may be necessary to work with lower fidelity versions of the clutter covariance model. One choice is,

$$\mathbf{R}_c = \sum_{p=1}^{P_c} \mathbf{v}(\theta_p, f_p)\mathbf{v}^H(\theta_p, f_p) \tag{11}$$

where we see that the scattered power for each clutter patch has been set to unity (for cases when no knowledge of the ground clutter reflectivity is available) and the matrix tapers $\mathbf{T}_p$ have been omitted (for cases when no knowledge of the clutter modulation is available). This matrix represents the ground clutter subspace and will require knowledge about the platform heading, speed, system PRF, relative antenna positions, and operating frequency. These are all parameters that are generally available in real-time. A computationally efficient method for computing (11) is given in [17]. This model will also require that the system be calibrated so that the spatial component of the response vectors $\mathbf{v}(\theta_p, f_p)$ can be computed. While a practical system will never be perfectly calibrated, we would expect that it will be calibrated to a level such that significant clutter cancellation based on the clutter subspace model in (11) will still result. For example, typical airborne radar systems achieve on the order of 20 dB worth of clutter cancellation using deterministic filtering techniques such as DPCA [18]. Finally, we note that when we use the clutter model given in (11) we would expect that a good choice for the colored loading level will be a value that results in the diagonal elements of $\mathbf{R}_c$ being approximately equal to an estimate of the clutter-to-noise ratio (CNR) on a single element and pulse. We note that an adequate estimate of CNR will typically be readily obtained by observing the radar data. By setting the loading level to a value that is close to the CNR we would expect that under conditions of perfect calibration the clutter model will null the deterministic ground clutter to the thermal noise floor, which is the desired result.

3.2. Data Domain Implementation

As discussed in Section 2, colored loading is a generalization of diagonal loading of covariance estimates whereby a general matrix, $\mathbf{Q}$, is added to the original covariance estimate $\mathbf{R}_{xx}$ to form a new color loaded covariance estimate

$$\mathbf{R}_{CL} = \mathbf{R}_{xx} + \mathbf{Q}. \tag{12}$$

An adaptive weight vector is then calculated using $\mathbf{w} = \alpha\mathbf{R}_{CL}^{-1}\mathbf{v}$, where $\mathbf{v}$ is the steering vector of interest. In [19], a computational assessment was performed for several approaches to computing this weight vector in both the covariance and data domain. In the former, the covariance estimate is explicitly computed and manipulated, while in the latter, computations are performed on the data directly, typically using the QR decomposition algorithm [20], without computation of the covariance estimate. In particular, a weight determination involving diagonal loading of the covariance estimate may be implemented in the data domain in a computationally efficient manner using the QR decomposition algorithm. This is accomplished by taking advantage of the structure of the diagonal loading matrix. In fact, the computational cost of this algorithm was shown to be essentially equivalent to that of simply computing the covariance matrix from the data

$$\mathbf{R}_{xx} = \mathbf{x}\mathbf{x}^H \tag{13}$$

where the appropriate normalization constant has been absorbed into the definition of the data vector, $\mathbf{x}$.

Colored loading may be implemented in the data domain for the same computational cost as diagonal loading if the color loaded matrix, $\mathbf{Q}$, has certain properties. Specifically, if the matrix $\mathbf{Q}$ is Hermitian and positive-definite, a new data vector, $\mathbf{x}'$, may be defined so that the desired colored loading matrix, (12), results

$$\mathbf{R}_{CL} = \mathbf{x}'\mathbf{x}'^H = \begin{bmatrix}\mathbf{x} & \mathbf{C}\end{bmatrix}\begin{bmatrix}\mathbf{x}^H \\ \mathbf{C}^H\end{bmatrix} = \mathbf{R}_{xx} + \mathbf{Q} \tag{14}$$

where $\mathbf{Q} = \mathbf{C}\mathbf{C}^H$ is the Cholesky decomposition of $\mathbf{Q}$ (which only exists if $\mathbf{Q}$ is Hermitian and positive-definite). The decomposition matrix $\mathbf{C}$ has the property of being lower triangular (i.e. all elements above the diagonal are zero). This extended data vector is very similar to that used in the diagonal loading implementation except that the augmented part there was a diagonal matrix (only diagonal elements were non-zero) whereas here the augmented part has non-zero elements below the diagonal as well. However, the QR decomposition, which is performed on $\mathbf{x}'^H$, only requires that the elements *below* the diagonal of the augmented part be non-zero to achieve the indicated computational efficiency. The augmented part in this case is C^H, which is upper traingular, and so satisfies this requirement. Thus the same computational cost as with diagonal loading results. Specifically, the cost of computing and applying the weights using the data domain representation (with colored loading) has about the same computational complexity as simply computing the covariance estimate, without loading, as in (13). Of course, in the covariance domain, the weight vector still needs to be computed, requiring an effective inversion of the loaded covariance, thus necessarily requiring more total computation than the the data domain representation described here. Furthermore, computation in the data domain is better conditioned numerically than in the covariance domain, which increases the range over which values vary, (13).

It is important to understand under what circumstances the color loaded matrix, $\mathbf{Q}$, satisfies the requirements of being Hermitian and positive-definite so that this approach may be used. A general class of colored loading matrices may be represented using

$$\mathbf{Q} = \beta_d \tilde{\mathbf{x}} \Lambda \tilde{\mathbf{x}}^H + \beta_L \mathbf{I}, \tag{15}$$

which possesses the necessary properties when the loading values, β_d, β_L, and the elements of the diagonal matrix, Λ, are real, positive quantities. This structure accomodates the low fidelity clutter covariance model suggested by (11), as well as higher fidelity models that include site-specific scattering weights and ICM tapers, as shown in (1) and [4,5]. It also accomodates reduced-DoF versions of these matrices, as obtained using the transformation process described by (6) and (7).

3.3. Two Alternative Implementation Approaches

We now consider two potential approaches to implementing reduced-DoF processing with knowledge-aided pre-filters in terms of their computational complexity. Approach #1 is to first apply the full-DoF knowledge-aided pre-filter to the data, and then apply the DoF-reducing transformation, followed by reduced-DoF adaptive beamforming. Approach #2, reverses the first two steps; specifically, first apply the DoF-reducing transformation to the data and then apply the reduced-DoF knowledge-aided pre-filter, followed by reduced-DoF adaptive beamforming. Note that although both of the preprocessing operations are linear, the whitening matrix for each of the two approaches is different, thus the two are not necessarily equivalent.

To compare these two approaches, we first represent them mathematically. Assuming that both the full-DoF loading matrix, $\mathbf{Q}$, (4), and its reduced-DoF counterpart, $\mathbf{Q}_m$, (10), satisfy the conditions outlined in the previous section, they can be written in terms of their respective Cholesky decompositions, $\mathbf{C}$ and $\tilde{\mathbf{C}}$, for full-DoF

$$\mathbf{Q} = \beta_d \mathbf{R}_c + \beta_L \mathbf{I} = \mathbf{Q}^{1/2}\mathbf{Q}^{1/2} = \mathbf{C}\mathbf{C}^H \tag{16}$$

and reduced-DoF

$$\mathbf{Q}_m = \beta_{d,m}\mathbf{R}_{c,m} + \beta_{L,m}\mathbf{I} = \mathbf{Q}_m^{1/2}\mathbf{Q}_m^{1/2} = \tilde{\mathbf{C}}\tilde{\mathbf{C}}^H \tag{17}$$

implementations respectively. For general $\mathbf{Q}$, the knowledge-aided pre-filter matrix is the inverse of its matrix square root, $\mathbf{Q}^{-1/2}$, a quantity that requires considerable computation to obtain. However, if $\mathbf{Q}$ is Hermitian and positive-definite, then the inverse of its Cholesky decomposition matrix may be used as the pre-filter instead. This not only requires less computation than the general pre-filter, but also permits the efficient data domain implementation described in the previous section. It should also be noted that it is not necessary to explicitly compute the inverse of the pre-filter matrices during implementation, as they are needed only as premultipliers of other quantities. Instead, techniques such as those described in [19] and the previous section are used. In fact, both approaches examined here may be implemented with the efficient methods previously described. However, for succint representation of these algorithms, the inverse notation is used.

With this notation, the resulting data vector with Approach #1, after the full-DoF pre-filter and then the DoF-reducing transformation, is

$$\hat{\mathbf{x}}_1 = \mathbf{H}_m^H \mathbf{C}^{-1} \mathbf{x} \equiv \mathbf{A}_1 \mathbf{x} . \tag{18}$$

The resulting data vector with Approach #2, after the reduced-DoF transformation and then a reduced-DoF pre-filter, is

$$\hat{\mathbf{x}}_2 = \tilde{\mathbf{C}}^{-1} \mathbf{H}_m^H \mathbf{x} \equiv \mathbf{A}_2 \mathbf{x} . \tag{19}$$

If the colored loading matrix is constant with range, then both sets of combined pre-filter/reduced-DoF matrices, $\mathbf{A}_1$ and $\mathbf{A}_2$, may be pre-computed once, adding only to the overhead of the entire operation. In this case, the combination that results in better detection performance should be used.

However, if the colored loading matrix is a function of range, then its respective pre-filter (i.e. Cholesky decomposition) must be re-computed every time its dominant subspace is adjusted. The computational complexity of this calculation scales as the number of DOFs to the third power so that it is clearly more expensive for the full-DoF pre-filter (Approach #1) than the reduced-DoF pre-filter (Approach #2).

Furthermore, as shown in [2], explicit pre-filtering results in diagonal loading of the pre-filtered covariance estimate, which is functionally equivalent to colored loading of the covariance estimate from original data. However, since the computation of the weights has the same computational complexity whether diagonal or colored loading is used in the data domain (when $\mathbf{Q}$ is Hermitian and positive-definite) and a pre-multiplication of one matrix by another equally sized matrix (i.e. pre-filtering) requires more computation than simply adding two matrices of the same size (i.e. colored loading) the computational complexity of Approach #2 may be further reduced with the use of colored loading rather than explicit pre-filtering. If full-DoF colored loading is used instead of full-DoF pre-filtering as the first stage of Approach #1, a Cholesky decomposition will still be required on the result of the reduced-DoF transformation since, in general, the application of this transformation matrix to the Cholesky decomposition of the full-DoF colored loaded data matrix will not preserve its desired triangular form.

Thus, we conclude that Approach #2 appears computationally less expensive than Approach #1. As a result, we have pursued the second approach in the results that follow. In the future, we will also compare the detection performance of the two approaches.

4. RESULTS

This paper is primarily addressing the problem of minimizing sample support requirements in a heterogeneous clutter environment (as opposed to the corruption of training data). Thus, the colored loading beamformer based on the lower fidelity clutter model given in (11) was applied to the KASSPER Workshop ‘02 *clutter-only* data set [21]. This data set simulates an L-band radar with parameters similar to the system used under the Multi-Channel Adaptive Radar Measurement (MCARM) program [22] and includes site-specific clutter computed using DTED Level 1. Therefore this data set represents a generally heterogeneous clutter environment. Moreover, array errors on the order of 5-10 degrees are included, see [21] for details. The simulated system has 32 pulses and 11 spatial channels, e.g. 352 DoFs. Figure 1 shows the eigenvalues for a representative range bin. We see that the effective clutter rank is approximately 50 (noise floor is at 0 dB on the plot).

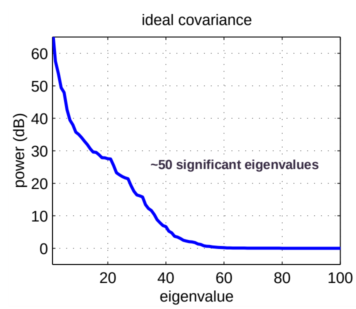

Fig. 1. Eigenvalues for a single range bin in the simulated data set.

Figure 2 shows the signal-to-interference plus noise ratio (SINR) loss [13] surfaces for a sample matrix inverse multi-bin post-Doppler element space reduced-DoF STAP algorithm. Three post-Doppler bins are used, $K = 3$, so that the number of reduced DoFs is $D = 33$. Results with 33, 66 and 99 range bins of sample support to estimate the covariance matrix are presented, with 0 dB of diagonal loading relative to the white thermal noise (for “diagonal loading-only”) and with the loading parameters β_L and β_d set to 0 dB and 30 dB relative to the thermal noise level for the colored loading beamformer. We see that the colored loading beamformer performance degrades much more gracefully as the sample support is decreased. This is a desirable property when one considers that operation in highly non-stationary interference environments (e.g., bistatics, severe terrain, etc.) often requires highly localized training regions.

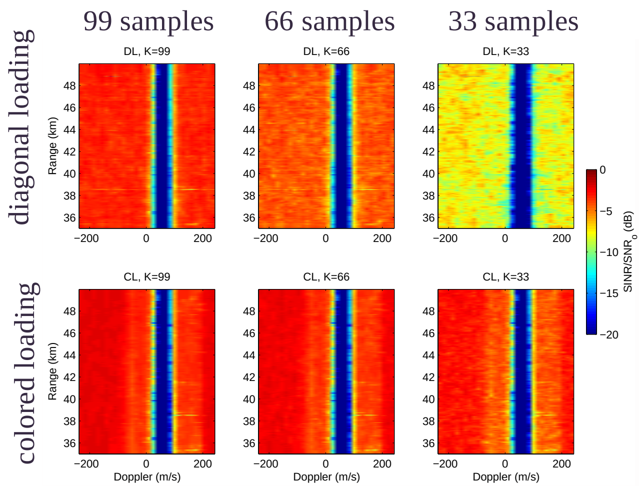

Fig. 2. Post-Doppler SINR loss surfaces for varying levels of sample support, for both diagonal loading only (top) and colored loading (bottom).

The space-time beamformed clutter data was post processed with a median constant false alarm (CFAR) algorithm that employed 100 range pixels and a single Doppler pixel to compute a background noise level for the beamformer output. The normalized beamformer output was thresholded and the number of false alarms was recorded. The sample processing steps were then applied to a series of test targets in all range bins for a specified Doppler shift. The target SNR was set to 25 dB at the closest range bin which represents a radar cross section of approximately 5 dBsm. The target SNR was then reduced versus range using an inverse range to the fourth power rule which resulted in a target SNR of 19 dB at the longest range in the simulation. The number of targets exceeding the thresholds when processed using the same beamformer weights and CFAR normalization was also recorded. P_d vs. P_{fa} curves were then generated using the observed number of false alarms and detections.

Figure 3 shows the P_d vs. P_{fa} curves for a Doppler bin that is very close to the clutter ridge (two bins away from the mainlobe clutter bin). We see that the colored loading beamformer performance degrades much more gracefully as a function of sample support than the standard post-Doppler SMI STAP algorithm with diagonal loading only. Also shown for comparison purposes are the results for a beamformer based on the ideal covariance matrix ("ideal cov.", with perfect knowledge of the array manifold), and a deterministic-only filtering derived using the colored loading matrix ("model-only", i.e. no adaptive component, $\mathbf{R}_{CL} = \mathbf{Q}$, and no knowledge of array errors). The reduced-DoF blended color loading approach clearly suffers some loss at this critical Doppler bin, relative to optimal processing, but is a considerable improvement over the deterministic-only approach. While not shown here, [2] presented a similar comparison of diagonal -only and colored loading results using full-DoF processing, with similar conclusions. There is an expected small performance loss with the reduced-DoF results relative to full-DoF, which should be balanced with the considerable reduction in computational complexity with the former relative to the latter, since the computation of the effective inverse of the covariance estimate needed for weight computation scales as number of DoFs to the third power.

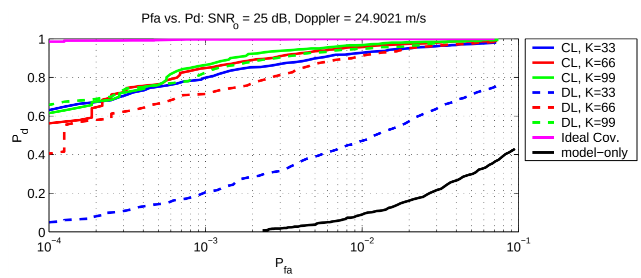

Fig. 3. KASSPER Simulation P_d vs. P_{fa} for a Doppler bin that is very close to the clutter

Figure 4 shows the same result for a Doppler bin that is well separated from the ground clutter. In this case we see that all of the algorithms with adaptivity perform well, although the diagonal-loading-only beamformer with the lowest value of sample support still results in somewhat degraded performance relative to the other beamformers. Thus, it is possible to achieve near-optimal performance with relatively low sample support (roughly equal to the number of reduced DoFs) and the colored loading beamformer under such conditions.

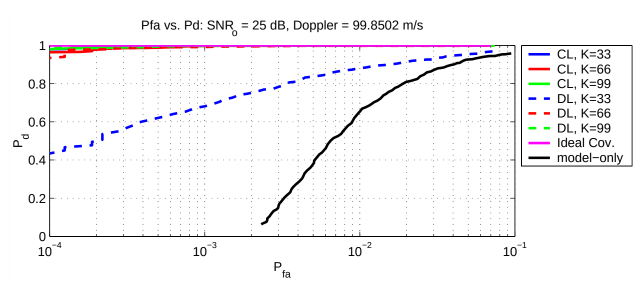

Fig. 4. KASSPER Simulation P_d vs. P_{fa} for a Doppler bin well-separated from clutter.

5. CONCLUSIONS

An approach to space-time beamforming that allows for inclusion of *a priori* knowledge about the ground clutter environment was presented and extended to reduced-DoF STAP implementations. The approach is based on constrained space-time beamforming and requires an *a priori* ground clutter covariance model. “Soft” quadratic constraints [1] were used that results in “colored” loading of the adaptive covariance estimate that can be implemented efficiently in the data domain. This approach offers a “blending” between adaptive and deterministic filtering. Practical implementations that do not require significant increases in available knowledge sources or computational resources were presented and analyzed. A preferred implementation, in terms of computational efficiency, was applied to a high fidelity simulated data set.

The techniques were shown to result in detection performance that is more robust, relative to adaptive processing only, to the limited sample support that may be prevalent when operating in highly non-stationary clutter environments, as well as better performance near the mainbeam clutter leading to improved minimum detectable velocity (MDV). Future work will focus on further demonstrating the utility of the approach on other heterogeneous data sets, both simulated and experimental, as well as its sensitivity to errors both in the *a priori* knowledge sources and the radar sensor (e.g. antenna calibration errors) itself.

6. REFERENCES

[1] B. D. Van Veen, “Minimum variance beamforming with soft response constraints,” *IEEE Trans. on Signal Processing*, vol. 19, pp. 1964-1972, September, 1991.

[2] J. S. Bergin,C. M. Teixeira, P. M. Techau, J. R. Guerci, “Space-time beamforming with knowledge-aided constraints,” *Proc. of the Adaptive Sensor Array Processing Workshop, MIT Lincoln Laboratory*, Lexington, MA, March 11-13, 2003.

[3] R. C. DiPietro, “Extended factored space-time processing for airborne radar systems,” *Record of the Twenty-Sixth Asilomar Conf. on Signals, Systems and Computer*s, vol. 1, pp. 425-430, October 26-28, 1992.

[4] P. M. Techau, J. R. Guerci, T. H. Slocumb, and L. J. Griffiths, “Performance bounds for hot and cold clutter mitigation,” *IEEE Trans. on Aerospace and Electronic Systems*, vol. 35, pp. 1253-1265, October, 1999.

[5] P. M. Techau, J. S. Bergin, and J. R. Guerci, “Effects of internal clutter motion on STAP in a heterogeneous environment”, *Proc. 2001 IEEE Radar Conf.*, Atlanta, GA, pp. 204 -209, 1-3 May 2001.

[6] J. S. Bergin, P. M. Techau, W. L. Melvin, and J. R. Guerci, “GMTI STAP in target-rich environments: site-specific analysis,” *Proc. of the 2002 IEEE Radar Conf.*, Long Beach, CA, 22-25, April 2002.

[7] W. L. Melvin and J. R. Guerci, “Adaptive detection in dense target environments,” *Proc. 2001 IEEE Radar Conf.*, Atlanta, GA, pp. 187-192, 1-3 May 2001.

[8] J. R. Guerci, “Theory and application of covariance matrix tapers for robust adaptive beamforming,” *IEEE Trans. on Signal Processing*, vol. 47, pp. 997-985, April, 1999.

[9] J. McClellan, G. Showman, and A. Saidi, “Direct clutter cancellation for STAP,” *Proc. of the 1996 Adaptive Sensor Array Processing Workshop*, Lexington, MA, 13-15, March 1996.

[10] W. L. Melvin *et al.*, “KA-STAP Development: GTRI Update,” Briefing presented at the KASSPER 2002 Program Review, September 23, 2002, MIT Lincoln Laboratory, Lexington, MA.

[11] H. Cox, *et al.* “Robust adaptive beamforming,” *IEEE Trans. of Acoustics, Speech, and Signal Processing*, vol. 35, pp. 1365-1376, October, 1987.

[12] K. L. Bell and K. E. Wage, “Partially adaptive LCMV beamforming with quadratic pattern constraints,” *4th. World Multiconf. on Systemics, Cybernetics and Informatics (SCI 2000)*, Orlando, FL, vol. VI, pp. 213-218, July 2000.

[13] J. Ward, “Space-time adaptive processing for airborne radar,” Lincoln Laboratory, Tech. Report 1015, December, 1994.

[14] J. S. Bergin, “A multi-stage filter structure for mitigating interference sources with partially known statistics,” ISL Tech. Note, ISL-TN-01-002, March, 2001.

[15] E. J. Baranoski, private communication.

[16] J. D. Hiemstra, “Colored diagonal loading,” *Proc. of the 2002 IEEE Radar Conf.*, Long Beach, CA, 22-25, April 2002.

[17] C. M. Teixeira “Closed form analytical expression for the clutter ring covariance,” ISL Tech. Note ISL-TN-02-005, July, 2002

[18] M. Skolnik, *Radar Handbook*, McGraw Hill, Boston, MA, 1990

[19] C. M. Teixeira and J. S. Bergin, “Comparison of algorithms for solving the covariance inverse problem for signal processing applications,” ISL Tech. Note ISL-TN-02-004, January, 2002

[20] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd Edition, Johns Hopkins University Press, Baltimore, MD, 1996.

[21] J. S. Bergin and P. M. Techau, “High fidelity site-specific radar simulation: KASSPER ‘02 Workshop datacube,” ISL Tech. Report ISL-SCRD-TR-02-105, May, 2002.

[22] D. K. Fenner and W.F. Hoover, “Test results of a space-time adaptive processing system for airborne early warning radar,” *Proc. of the IEEE Radar Conf.*, Ann Arbor, MI, pp. 88-93, 13-15, May 1996.